Bar,
ABA.8
E4m

MULL

Modern Uses of Logic in Law: Quarterly Newsletter of the Electronic Data Retrieval Committee of the American Bar Association. Electronic...

60J

Contents 1960 June

		Page
1. APPLICATIONS AND REPORTS ON RESEARCH		
a. Information Retrieval Research in the U.S. Patent Office	Simon M. Newman	45
b. The Industrial Revolution in Law	Lee Loevinger	56
c. The MINICARD System for High Speed Data Retrieval	Allen M. Rahm	63
d. Storage and Retrieval of Legal Information: Possibilities of Automation	Layman E. Allen, Robin B.S. Brooks & Patricia A. James	68
2. LOGICAL THEORY		
3. REVIEWS AND ABSTRACTS		
a. Machine Data Processing Systems for the Trial Lawyer	Roy N. Freed	85
4. BIBLIOGRAPHY		
a. Bibliography on Data Retrieval for Legal Research	Richard F.C. Hayden	86
b. Selected Bibliography of Sources of Current Information on Documentation		87
5. PROBLEMS		
6. READERS' LETTERS AND OTHER MISCELLANEOUS ITEMS		
a. Policy Statement of West Publishing Company on Automatic Retrieval of Legal Information	William W. Marvin	89
b. Program of the EDR Committee at the 1960 Annual Meeting of the ABA in Washington, D.C.	Reed C. Lawlor	90
c. Practicing Law Institute Course Dealing with Information Retrieval		92

UNIV. OF MICH.
AUG 29 1960
LAW LIBRARY

For ease of **Information Retrieval**

The AM JUR-
ANNOTATED REPORTS SYSTEM

The income-building library that brilliantly solves modern-day law practice problems

The interlocking law library that gives you in one logical arrangement: the law in breadth and in depth, almost every conceivable legal or business form, and pleading and practice form, and the way to prove your facts in court. Write today for full details.

The Lawyers Co-operative Publishing Co.
Rochester 3, New York

MULL is published quarterly (Mar., June, Sept. and Dec.) by the Electronic Data Retrieval Committee of the American Bar Association. Edited by Layman E. Allen with the advice and assistance of Alan R. Anderson, Mary E. Caldwell, Patricia A. James, Harold D. Lasswell and Leon Lipson. Send manuscripts to: MULL 401A Yale Station, New Haven, Conn. Subscription rates: $1 for single copies, $2 for one year, $5 for three years.

The views expressed in MULL are to be attributed to their authors and not to MULL, nor to its editors nor to the Electronic Data Retrieval Committee.

1. APPLICATIONS AND REPORTS ON RESEARCH

a. INFORMATION RETRIEVAL RESEARCH IN THE U.S. PATENT OFFICE*

Introduction

In its mission of issuing patents for inventions, the U. S. Patent Office is required by law (R1) to reject applications for patents which do not, inter alia, claim novel subject matter. The Office has a professional staff of scientists, each of whom examines the applications in a small sector of science and technology (the so called prior art). The first step in the examination is a search of the prior art to locate any document, or combination of documents, which discloses any subject matter from which the examiner can evaluate the question of novelty presented by the application. The documents so obtained are known as references. Of course, the other requirements in addition to novelty which govern issuance are also considered during examination. The examining corps spends approximately 60% of its time merely searching (as opposed to the remaining steps, including evaluating the retrieved references) in what is one of the largest information retrieval operations in existence.

In 1954, in its consideration of a Department of Commerce budget request, the U. S. Senate Appropriations Committee (R2) directed that the Department

> "... make an aggressive and thorough investigation as to the possibility of mechanizing the searching operations ... in order to modernize, insofar as practical, the Patent Office operations."

The Secretary of Commerce appointed an Advisory Committee, which recommended in part (R3):

> "... A Research and Development unit should be established in the Patent Office.... The National Bureau of Standards and the Patent Office should undertake a joint program to stimulate and develop ... techniques specifically adapted to Patent Office operations...."

On January 31, 1955, a small unit was formed within the Classification Operations of the Patent Office, and on April 8,

---

*The content of this article is in the public domain.

1956, this matured into the Office of Research and Development, directly under the Office of the Commissioner of Patents. On August 3, 1957, Mechanized Examining Division A was created under the Office of Research and Development to handle the automated search of the first sector of the prior art which had been prepared for search mechanization.

Our Problems of Retrieval

It has been stated many times that the very discriminating search which the Patent Office makes is different from the types of searches made by others for information. It is our studied opinion that this is not true, that our searching and other searching differ merely in degree, and not in kind. Others have either been afraid to hope for such detailed retrieval, or have abandoned it as visionary. But we have been retrieving by manual methods with this degree of detail for years, and research scientists, who have taken the trouble to come to Washington to use our search files, have uniformly been surprised at the level of detail which has been provided.

Successful mechanization of a manual system does not necessarily consist in mechanically performing the same steps in precisely the same way. The mechanized horse and surrey is not a tractor-trailer unit, but an automobile; the domestic coak stoker feeds coal from the bottom of the furnace bed, many commercial coal burning furnaces pre-grind the coal, mix it with air and blow it into the firebox, while manually fired furnaces are stoked by adding coal on top of the fire bed. Successful mechanized retrieval systems may thus prove to be substantially different in operation from the conventional search systems which they supersede.

Patents and other documents written in expository prose, even when referring to the same phenomenon, will use terms on different levels of generality and specificity. An article, device, method, or other phenomenon may be described by any of a number of different facets or aspects of its nature. To further complicate this situation, such facets or aspects usually occur on different levels of generality. Furthermore, much significant information in technical prose is present by implication, rather than explicity expressed.

As a simple illustration of two levels of generality, a dog, cat, parakeet, and guppy are each "vertebrates", but only the

first two are "mammals". However, to some of us, a dog, cat, parakeet, and guppy may also belong to the genus "pet". Simultaneously, to others, the same dog may belong to the genus "burglar protection", together with a padlock and a policeman. This cat also belongs to the genus "rodent destroyers", together with a mousetrap and arsenic. This illustrated facets or aspects at two levels of generality, but with the same species subsumed from different genera.

Implicit disclosure is evidenced, for example, by specification of the liquid state of a compound which inherently suggests the existence of a freezing and a boiling point, a specific gravity, and other physical properties. If the compound is a dye, a color is implied; if a pharmaceutical, a disease or other bodily malfunctioning and a process of treatment or cure or both are implied; if an explosive, a method of and means for both priming and igniting; etc.

In the process of retrieving similar and analogous subject matter, the ability to specify interrelationships among units of subject matter (articles, devices, method steps, and other phenomena) enhances precision of the search and reduces the incidence of false or unwanted answers. The utility of interrelationship is seen in contrasting the information value of parts lists and circuit diagrams for television and radar sets. An expert electronics engineer, if shown only the two parts lists, would have difficulty in identifying the list with the appropriate device, but when presented with only the circuit diagrams even without values for the individual components, he could immediately distinguish between the two because of the interrelationships of the components shown.

The user of a retrieval system for any file of documents cannot, of course, know whether a document discloses or implies the desired subject matter, or, if present, the level of generality and the facets or aspects by which it has been described. To be of optimum value, therefore, a system should provide for each of these factors in its file preparation and its searching strategy.

Our Research Program

The Patent Office and the National Bureau of Standards, jointly, separately and with outside contractors (including the Bureau of the Census), have been and are still engaged in a series of programs which range from actual productive mechanized

searching of a limited group of chemical compounds to exploratory research in information retrieval theory including mathematics and linguistics. Most of the early research effort was in the field of chemical subject matter. The scope was later broadened to include studies of problems posed by language and its ambiguities and by disclosures in the mechanical and electrical arts. Our own studies have tended to emphasize problems having immediate application to the patent examining operation over those having projected applications of broader scope.

Our first attempt at automation of patent searching antedated the Congressional mandate. This research was done in 1947 in the Chemical Division of the Classification Group of the Patent Office. A mechanized search was actually attempted in 1947, but the proposed system was not sufficiently sophisticated for patent searching. Other systems were conceived and studied, and in 1950, an operative system was tested (R4). After the test, work was suspended for administrative reasons until 1955.

The reconstituted research and development group began with a reappraisal of this first attempt followed by work on systems design. Subsequently, several of the new group began a cooperative program with the Data Processing Systems Division of the National Bureau of Standards. This cooperative program, christened HAYSTAQ, is an experimental effort designed for long range studies to develop a completely comprehensive search routine for a large Data Processing Computer (R5). Although specifically tailored to chemical disclosures, general applicability of the system to any topological network is planned. The system has been designed to provide for searching for compounds, compositions of matter (i.e., mixtures of compounds), processes of manufacture and use, and other related matters. In the construction of the computer program, all of the desired limitations of a patent search have been considered, so that this program may later be extrapolated to eventually retrieve substantially all information in a publication. Currently, a comprehensive search file of chemical compounds in one segment of the chemical arts is being prepared for use in a first approximation of the desired type of search.

Another early effort (R6) was an analysis of the examiner's duties in patent examination with particular emphasis on the problems entailed in mechanization of search procedures.

By contrast with these early projects, the amenability of the structures of chemical compounds to routinized designations has

facilitated the development of several immediately usable systems for the small areas of subject matter to which they apply. Perhaps the most spectacular is the coding system for steroid compounds by which many variations of the present structure have been represented in generic and specific aspects (R7). Steroid chemistry constitutes one of the more complex organic chemistry arts. The coding system which was devised is a relatively unsophisticated method by which the steroid compounds content of the approximately 1600 U. S. patents in this art was entered into a series of punched cards with flexibility in generality sufficiently detailed for productive search. Machine search of the punched-card file as part of the examination of patent applications in this art commenced in September 1957, in the Mechanized Examining Division. From that time through May 1960, approximately 14,000 mechanical searches have been made by examiners in this division. This system has been widely acclaimed and publicized; one measure of its value is indicated by the undertaking of the Pharmaceutical Manufacturers Association to finance a literature supplement to the patent file. A method for extracting and preparing the steroid data for machine input was established which eliminates, as far as possible, manual repetition and transcription of data, thus conserving time, space, and professional effort.

In conjunction with early work on the steroids, a limited topological search for organic compounds, conceived in the Patent Office, was programmed by National Bureau of Standards personnel (R8) for their SEAC computer. While adequate to minimal needs for steroid compounds, it proved cumbersome and slow, and was almost immediately superseded by the present productive system.

A system, known as the VS_3 system (R9), developed for serially searching all known types of chemical structures, subsequently served as a springboard for several current projects. In this system, generic and subgeneric terminology are expressed in terms representative of chemical structure diagrams, i. e., a generic term is applied to a chemical structure if a substructure or "fragment", defined by that term, is a part of the structure so diagrammed. Each term is assigned a distinctive code, which is further modified by an additional coding, called an "interfix", representing diagrammed connectivity between the fragments.

This general scheme was later redesigned for use in a random access approach, in which the stored information is somewhat loosely called an "inverted file" (R10). Basic principles applied in this new approach include developing genus-species hierarchies

from the extracted raw data of the document, rather than by intuition or preconception; the ability to manipulate and modify the file without altering encoded data to changes in classification and relationship on the generic level; and, when utilized in a retrieval operation, constructing both the classification hierarchy and the variant and invariant relationships from the subject-matter file without dependence on document identification.

One of the descendants of this modified VS_3 system is project RAMP, in which the modified VS_3 approach has been adapted to compounds containing the phosphorus acid ester structure (R11). The complexities arising in this latter system from a more variant structure are manifest in greater generality of descriptors and interrelationships. This system is now in pilot operation, and a search file is being created for productive use.

Another descendant is project DRAM, in which the inverted file technique of the modified VS_3 system is being extended to include non-structural descriptive matter pertaining to chemical behavior and enviroment.

In a recently started project, electrical systems components and fragments of such systems have been identified with distinctive codes at several levels of generality. To these concepts, additional descriptive terms have been developed for functions and other identifiable characteristics.

Other means of standardizing terminology for retrieval which are being studied include a study of symbolic and graphical systems of representing word and diagrammatic disclosures of mechanical subject matter (R12).

The difficulties involved in being able to represent a concept in any technical discipline in unique, precise, and unambiguous terminology led to more basic research into linguistics. Although the amount of effort applied in this direction has been limited, many of the problems imposed by popular language usages have been stated (R13). Some suggestions on studies leading to their resolution have been proposed, and exemplary steps taken, including the formulation of one possible solution, a metalanguage, dubbed "Ruly English" (R14). Concomitantly, through contracts made by the National Bureau of Standards, research is going forward on using raw text as a file in information retrieval (R15). Creation of a file of U. S. Patents in machine-readable form for this research required a complex transliteration scheme (R16).

The file is now in preparation.

In conjunction with research on retrieval systems, some study has been made of hardware for processing the data. To test retrieval systems utilizing network representations, radical modifications were made of the Census type Multi-Column Sorter (similar to the IBM 101) in a joint effort with the Bureau of the Census of the Department of Commerce (R17). The resulting machine, named the ILAS (Interrelated Logic Accumulating Scanner), is programmed in part by setting subject matter codes and in part by designating interrelationships at different logical levels among the subject-matter codes. The built-in ability of the machine to make a variety of correlations is used for rapid testing of the logical design of coding systems. Existing equipment is also being used or has been evaluated for testing our retrieval systems. These include the National Bureau of Standards Computer SEAC for which programs are being written to test HAYSTAQ; the IBM RAMAC 305 computer which has been programmed to test "inverted-file" type arrangements of subject matter (R18); the Bendix G-15 computer which was programmed to evaluate the use of weighted values imposed on the subject matter codes (R19); and the "Peek-a-Boo" card (Batten-type card) equipment of the Office of Basic Instrumentation, National Bureau of Standards.

Future Directions

Several factors make our future course unclear, a situation probably shared by other research groups. Preliminarily, continued progress on the retrieval systems currently being devised may disclose areas of emphasis not presently contemplated, or need for other unconsidered modifications. Funding and staffing are also factors not without significance. Additionally, a systems-engineering study of our operation is being planned by the Department of Commerce as a result of a Department-wide evaluation of its role in science and technology research (R20).

The creation of productive but limited search systems in small areas of art, when well considered, may be helpful in extrapolating both the sophistication of the system and the size of the area in each succeedingly developed system. Our experience in obtaining small-scale files for pilot tests strengthens our conviction of the need of studies into methods for preparing large encoded document files, including abstracting methods, in which mechanization should be considered at all possible steps. Development of screening techniques for serial systems and other methods of shortening search

time also need consideration.

Concomitantly, other more basic studies seem necessary for the resolution of problems which separate us from a universal solution to the patent-search problem. Subject areas which might complement and carry forward present findings could profitably include mathematics and the formulation of a mathematical theory of retrieval; structural and transformational linguistics; programming studies for learning and self-organizing systems; the development of pattern-recognition machines capable of accepting letters, pictures, and other graphic material as inputs; and the study of the possible applicability of machine-translation results, including the mechanized searching of non-English language documents.

The best envisaged plan of research can break down at the most unexpected places. Additionally, unsoundly conceived or inadequately considered steps in a program will not invariably expose themselves early in subsequent development or use, and will require more time to reform than was "saved" in their establishment. The obviation is apparent -- sufficient funds to do the job, sufficient time to do it in, and sufficient skilled staff to do it properly.*

Simon M. Newman
Acting Associate Director
Office of Research and Development
U. S. Patent Office

References

R1 U. S. Code; Title 35, Patents; Chap. 12; Sect. 131.

R2 U. S. Congress. Senate Committee on Appropriations. Report No. 1541, Departments of State, Justice and Commerce Appropriation Bill, 1915: report to accompany HR 8067. Washington, U. S. Govt. Print. Off., 1954, p. 21. (82nd Cong., 2d sess.)

R3 Report to the Secretary of Commerce by the Advisory Committee

---

*The author wishes to express deep appreciation for the unstinting and continuing assistance of Mrs. Rowena W. Swanson, who made constructive suggestions for the improvement of the original draft, prepared the bibliography, and carefully edited and proofread the completed manuscript.

on Application of Machines to Patent Office Operations. Washington, U. S. Dept. of Commerce, 1954, p. 15, 16.

R4 Bailey, M. F., B. E. Lanham, and J. Leibowitz. "Mechanized Searching in the U. S. Patent Office." Journal of the Patent Office Society 35: 7 (Aug. 1953) 566-587.

R5 Pfeffer, Harold, Herbert R. Koller, and Ethel C. Marden. A First Approach to Patent Searching Procedures on Standards Electronic Automatic Computer (SEAC). Patent Office Research and Development Report No. 10. Washington, U. S. Dept. of Commerce, 1958. Also: Koller, Herbert R., Ethel Marden, and Harold Pfeffer. "The HAYSTAQ System: Past, Present, and Future," in Proceedings of the International Conference on Scientific Information, Washington, D. C., Nov. 16-21, 1958. Vol. 2. Washington, National Academy of Sciences - National Research Council, 1959, p. 1143-1179.

R6 Newman, Simon M. Problems in Mechanizing the Search in Examining Patent Applications. Patent Office Research and Development Report (No. 3). Washington, U. S. Washington, U. S. Dept. of Commerce, 1956, revised 1959.

R7 Frome, Julius, and Jacob Leibowitz. A Manual for Coding Steroids. Patent Office Research and Development Report No. 11. Washington, U. S. Dept. of Commerce, 1958.

R8 Ray, Louis C. and Russell A. Kirsch. "Finding Chemical Records by Digital Computers". Science 126:3278 (Oct. 1957) 814-819.

R9 Leibowitz, Jacob, Julius Frome, and Don D. Andrews. "Variable Scope Search System: VS_3", in Proceedings of the International Conference on Scientific Information, Washington, D. C., Nov. 16-21, 1958. Vol. 2. Washington, National Academy of Sciences - National Research Council, 1959, p. 1117-1142.

R10 The term "inverted file" has historically been applied to manual multiple-entry systems which provided for segregation of only that portion of the file containing information pertinent to the subject matter of the search question. This arrangement is contrasted with the serial arrangement in which a sequential search of the entire file is necessary. See: Metcalfe, John. Information Indexing and Subject Cataloging - Alphabetical: Classified: Coordinate: Mechanical. New York, The Scarecrow Press, Inc., 1957, p. 171-174.

R11 Frome, Julius. Semi-Automatic Indexing and Encoding. Patent Office Research and Development Report No. 17. Washington U. S. Dept. of Commerce, 1959.

R12 Smith, Charles G. "Descriptive Documentation," in Proceedings of the International Conference on Scientific Information, Washington, D. C., Nov. 16-21, 1958. Vol. 2. Washington, National Academy of Sciences - National Research Council, 1959, p. 1097-1116.

R13 Newman, Simon M. Linguistic Problems in Mechanization of Patent Searching. Patent Office Research and Development Report No. 9. Washington, U. S. Dept. of Commerce, 1957.

R14 Andrews, Don D., and Simon M. Newman. Storage and Retrieval of Contents of Technical Literature - Nonchemical Information - Preliminary Report. Patent Office Research and Development Report (No. 1). Washington, U. S. Dept. of Commerce, 1956.

Newman, Simon M. Storage and Retrieval of Contents of Technical Literature - Nonchemical Information - First Supplementary Report. Patent Office Research and Development Report No. 4. Washington, U. S. Dept. of Commerce, 1957.

Newman, Simon M. Storage and Retrieval of Contents of Technical Literature - Nonchemical Information - Second Supplementary Report. Patent Office Research and Development Report No. 12. Washington, U. S. Dept. of Commerce, 1958.

Newman, Simon M. Analysis of Prepositionals for Interrelational Concepts - Preliminary Study. Patent Office Research and Development Report No. 16. Washington, U. S. Dept. of Commerce, 1959.

R15 Yngve, Victor H. "Feasibility of Machine Searching of English Texts," in Proceedings of the International Conference on Scientific Information, Washington, D. C., Nov. 16-21, 1958. Vol. 2. Washington, National Academy of Sciences - National Research Council, 1959, p. 975-995. Also: Yngve, Victor H. "In Defense of English". Preprints for An International Conference for Standards on a Common Language for Machine Searching and Translation, Cleveland, Ohio, Sept. 6-12, 1959. (Proceedings being published by Interscience Publishers, Inc., New York).

R16 Newman, Simon M., Rovena W. Swanson, and Kenneth Knowlton.

A Notation System for Transliterating Technical and Scientific Texts for use in Data Processing Systems. Patent Office Research and Development Report No. 15. Washington, U. S. Dept. of Commerce, 1959. (Revised paper being published in Proceedings of An International Conference for Standards on a Common Language for Machine Searching and Translation, Cleveland, Ohio, Sept. 6-12, 1959, by Interscience Publishers, Inc., New York.)

R17 Andrews, Don D. Interrelated Logic Accumulating Scanner (ILAS). Patent Office Research and Development Report No. 6. Washington, U. S. Dept. of Commerce, 1956. Also: Frome, Julius, H. R. Koller, Jacob Leibowitz, H. Pfeffer, and Don D. Andrews. Recent Advances in Patent Office Searching: Steroid Compounds and ILAS. Patent Office Research and Development Report No. 8. Washington, U. S. Dept. of Commerce, 1957. (Reprinted as Chap. 25 in Advances in Documentation and Library Science. Vol. 2. Information Systems in Documentation. New York, Interscience Publishers, Inc., 1957.

R18 Leibowitz, Jacob, Julius Frome, and Don D. Andrews. Variable Scope Patent Searching by an Inverted File Technique. Op. cit.

R19 Frome, Julius, Jacob Liebowitz, Don D. Andrews, Joseph D. Grandine, Steven T. Polyak, and Karl G. Siedschlag, Jr. A System of Retrieval - Compounds, Compositions, Processes, and Polymers. Patent Office Research and Development Report No. 13. Washington, U. S. Dept. of Commerce, 1958.

R20 The Role of the Department of Commerce in Science and Technology: A Report to the Secretary of Commerce by a Special Advisory Committee of the National Academy of Sciences. Washington, National Academy of Sciences - National Research Council, 1960.

b. THE INDUSTRIAL REVOLUTION IN LAW

There is no man living in a modern civilized country whose life has not been deeply affected by the scientific and industrial revolution of the last two centuries. We eat pasteurized milk and frozen meats and vegetables, dress in clothes made of synthetic materials, listen to radio, watch television, converse on the telephone, travel by automobile and airplane, sleep on foam rubber mattresses beneath electric blankets, and generally are so surrounded by the new products of science and industry that last year's miracles are today's banalities. The great majority of our population works at jobs that were created by scientific and industrial changes during the last century, whether the individual function is in production, distribution or administration. Now these jobs are being altered or abolished by automation, and a new category of technical occupations is arising to provide service for the automatons.

The professions, too, have generally changed in many significant aspects in response to the same influences. Today there are fundamental differences from twenty-five years ago in the practice of accounting, agriculture, architecture, engineering, journalism, and, perhaps above all, medicine and the health fields. There is one eminent and lonely exception--law. For all the multiplication of precedents, statutes and regulations, there has been no significant adaptation of legal techniques or practices to any of the new developments in science, technology or industry. Lawyers today are doing the same things in the same way that lawyers did a quarter, half and full century ago. Now we face the question whether this condition will continue.

What will the practice of law be like a quarter of a century from now?

So far as present knowledge supports a prediction, it will continue to be the same in many important respects as it is today. But it is even more certain that in many other significant aspects it will be quite different. Of course, the lawyer and his practice will be affected, both directly and indirectly by the innumerable changes that will occur in the community pattern of living as the result of scientific and technological developments. For example, it seems probable that typewritten letters may be transmitted across the country by electronic facsimile reproduction at least as rapidly as telegrams are today.

Of much greater significance to the legal profession is likely to be the revolution in the methods of storing and retrieving information, and the consequent changes in legal research and the daily work routine of the lawyer and judge. Based upon presently known techniques and devices now deing utilized in other fields, the following sketch of legal research and practice in 1985 can be projected as a conservative prediction.

The lawyer's task of research in 1985, as in 1960, will start with his own analysis of the problems into legal issues and categories. With some ideas of the categories in which answers to his questions may be found, the lawyer will turn to a one volume outline of key numbers that will, or should, be on every lawyer's desk constantly. This volume will be large but not too unwieldy, approximately a single volume of the present American Digest or ALR series in size. This volume will have an outline, similar no doubt to the outline now contained in the American Digest system or Corpus Juris Secondum, but somewhat more comprehensive. However, the outline will contain only topical headings and a system of code numbers. There will be no entries under any of the topical headings, and thus the outline itself will not become outdated rapidly as the present digest volumes do.

The lawyer confronted with a legal problem will consult his digest volume for all the topical headings that he thinks are or might be relevant or helpful to his search for legal precedents or rules. From each digest heading he will secure a code key number which he will write down on a legal data requisition card, printed in an appropriate form for this purpose.

After all the digest topic headings have been transferred to the data requisition card, the lawyer will consult a table printed on the inside front cover of the digest volume for the appropriate code symbol to indicate the kind of legal data he wants. By these symbols the lawyer will indicate whether he wants data under the digest headings relating to his subject matter for his own state jurisdiction alone, other specified state jurisdictions, or all of them, or for the federal jurisdiction, and also whether he wants court decisions, statutes, administrative regulations, administrative precedents, or any combination of some or all of these.

How long it will take to fill out such a card will depend upon the complexity of the problem and the number of subject headings involved, the facility of the individual, and the familiarity of the individual lawyer with the particular

subject matter. In many cases no more than a few minutes may well suffice. In others, it may take somewhat longer, but hardly ever more than an hour at most in even the most complex and difficult matter.

When the data requisition card has been completed the lawyer will give it to a messenger or an assistant, if he is in a hurry, or mail it to the local law library if he has plenty of time or resides out of town. The law library will consist primarily of several banks of electronic data handling machines. The librarian will take the data requisition card and transfer the code symbols to a punch card on a small typewriter-like machine at her desk. (Larger law offices may have such a machine in the office and prepare their own punch cards to save time.) The punch card will then be inserted into the control slot of the electronic "Law Digest Machine", a switch will be flipped, and there will be a slight hum as a number of printed sheets slide into a receptacle on the front of the machine. Each sheet will contain a photographic reproduction of the digest of each case, statute, regulation, administrative precedent or other relevant data indicated as desired on the data requisition card which has been classified to any of the subject headings listed on the card. It should take less than two minutes to make the punch card, insert it in the machine and secure copies of digests of all relevant items.

The machine itself will be merely an electronic digital computer, much similar to numerous such machines in general use in many industrial and scientific applications. It will have a control panel looking something like the console of an organ. Behind this will be a cabinet perhaps ten feet wide and ten feet high. The front several feet of the cabinet will contain an immensely intricate maze and mass of switches, relays, transistors, diodes, and printed and solid state electrical circuits. This is the "analyzing" or "thinking" part of the machine. It will have no moving parts as all "action" will be electrical. This part of the machine will "read" the message on the punch card and actuate circuits corresponding to each subject heading and category of data sought.

Behind the control panel and the electronic "brain" of the machine, there will be a bank of reels, mounted side by side and one above the other, and controlled by tiny electrical motors. Each reel will be about half an inch wide and about two inches in diameter. It will contain many feet of an extremely thin plastic tape which has a special coating on it to provide a thermoplastic record. Data is imprinted on the tape by writing with an

electrobeam that sweeps back and forth across the tape when recording is done, leaving a groove similar to, but much finer than, the groove on a long-playing phonograph record. By electrical means, this groove is melted and then frozen into the tape, making a lasting record. Each tape is capable of recording and storing the equivalent of several hundred volumes of ordinary printing. In a bank no more than half a foot in thickness and ten feet square there are about six thousand of these reels. Each six inch thick bank is the equivalent of a library of over a million volumes in stored information!

As the electrical circuits actuate the reels and their controlling mechanism, certain reels containing the desired data turn to pass particular portions of their tapes over a "scanner". The scanner "reads" the messages printed on the exposed parts of the tapes and transmits these to an image-reproducer which is much like an ordinary television tube. The fourth major component of the Law Digest Machine is the reproduction segment. Here the image-reproducer lights up brightly for a fraction of a second with a reproduction of the printed message recorded on the exposed portion of the tape. As this occurs, a photographic sheet is exposed and developed, and then dropped into the "out basket" on the control panel. It will be printed in a normal type face and be at least as legible as a good photostatic reproduction of 1960.

Thus in the time it takes the messenger to get a drink of water from the old-fashioned cooler in the corner of the room, the machine will have prepared printed copies of all digest paragraphs that have been subsumed under any of the headings and categories in which the lawyer indicated an interest. These will be returned to the lawyer who can then read them at his convenience.

Some of the digests will certainly appear to be irrelevant and will be discarded. Others will probably be saved for future reference in the file. Some will certainly require further investigation. These may refer to cases, statutes or regulations that are available in the lawyer's library of conventional books. However, as the number of decisions, statutes, regulations and rulings continued to multiply it will become increasingly difficult and expensive and commensurately less useful or profitable to attempt to maintain a reasonably current law library on the basis of conventional volumes. Consequently law office libraries everywhere will become smaller, rather than larger, as more and more reliance comes to be placed upon the central library with

its immense and efficient machines for the electronic retrieval of data. Therefore, it is likely that in many situations the digests will indicate important material to be examined in full that is not at hand in the office law library. But this will not involve any difficulty.

Each digest will itself bear a code designation. The lawyer will note on another data requisition card the numbers of each digest referring to material he wants in full text. The messenger will take the second card back to the law library, and in a similar brief period from essentially the same kind of machines, will obtain the complete text of each desired case, statute, regulation or other reference. The lawyer who desires to do so can, of course, take his own requisition cards to the law library; examine the digests immediately; secure the texts of such references as he desires at the time; and then return to his office for further study and analysis of the authorities thus secured.

This system offers many obvious advantages over the cumbersome and inadequate system of legal research employed in 1960. Perhaps the most important advantage is that it offers the lawyer - for the first time in modern history - the opportunity to do a complete job of legal research on any given problem within whatever limits may be specified as relevant. The vast capacity of the electronic machines for storing and retrieving data will permit classification, storage and retrieval in a systematic manner not only of appellate precedents, but also of statutes, administrative regulations and administrative and executive rulings. Until electronic retrieval of data is a reality, ignorance of at least administrative law will inevitably continue to be more the rule than the exception.

Another advantage of electronic research will be the assurance that the results are up-to-date. There is no field of law in which the 1960 system of reports, digests, supplements and loose-leaf reporters gives real assurance that all of the recent cases, statutes and regulations will be noted, even with extensive and arduous cross-checking. However, once a reference is reported and recorded in the Law Digest Machine, it will automatically come out in response to an input of the appropriate digest or code number. The most recent digests will be as quickly and easily accessible as the most ancient.

Aside from the assurance it gives to lawyers of completeness and efficiency, the greatest advantage of the electronic retrieval system of legal research will be its ability to save

a lawyer's time. Instead of looking through numerous volumes and supplements in order to secure authorities on a point, the search will be confined to a single, familiar outline digest that gives the code numbers needed to actuate the machines. From this point on, references are secured merely by pushing figurative electronic buttons for the digest numbers. Eventually, of course, the references must be read and analyzed; but the time spent in merely finding potentially relevant authorities will be reduced from hours to minutes.

As incidental advantages will be the convenience of having authorities furnished in separate, convenient, printed form. Additional copies will be equally available. Such copies may thus be marked as desired and filed with other papers without the necessity of taking valuable secretarial time to make length extracts from books.

For the small town lawyer, the electronic retrieval system will be a real equalizer. Under the 1960 system a fairly extensive law library is needed for thorough research of any legal point. Such libraries are usually available only in the larger metropolitan centers. On the other hand, long distance research through a correspondent is difficult, expensive and usually unreliable. But under the electronic retrieval system, the only difference between the lawyer most remote from the largest city and one living in it will be the time it takes for material to be transmitted by mail. The small town lawyer will have the same digest outline of key numbers as the metropolitan lawyer. He will use the same data requisition cards. And his cards will produce the same data from the same machines.

The one advantage that the electronic retrieval system will not have is economy. The completeness, efficiency, speed, convenience and availability of the system will have to be paid for. Undoubtedly payment will be made by those who use the system. The most obvious method of charging and collecting payment is by exacting a small fee for every page of data that is delivered on any lawyer's request. The charge could be as little as a fraction of a cent or, perhaps, as high as ten cents a page. It might well vary with the type of material sought, being different for digest pages, statutes and decisions. Presumably some arrangement for monthly deposit or charge accounts would be worked out, so that most lawyers would not have to keep a handful of petty cash around for every trip to the law library. It is impossible to estimate whether the aggregate cost of electronic legal research will be more or less than the cost of maintaining a rea-

sonably useful office law library in 1960. However, it does seem clear that taking time and efficiency into account, the more modern system will be far more profitable for the lawyer - and his clients - than the cumbersome and inefficient hunt and check system presently in use.

One aspect of legal practice that will not be changed merely by instituting a system for electronic retrieval of legal data is the thinking of lawyers. The revolution that we have been contemplating is an industrial revolution--a change in techniques and patterns of work resulting from advances in thinking in other fields. This is not a scientific revolution--a change in the basic thought patterns in our own field. On the contrary, any system of electronic retrieval of data will necessarily be based upon a classification of data made before the system itself is put into operation. Thus it will have some tendency to carry forward whatever categories and cliches inhere in the thinking of lawyers when the system is established. This does not mean that the law will become fixed and immutable--any more than the system of digests and reports has made the law fixed and immutable. Indeed, it may well be that electronically recorded data can be more easily reclassified than that which is preserved in the conventional printed page. But there is nothing inherent in the electronic retrieval of data itself which will alter any legal concepts or procedures or make the legal system any more flexible or responsive in substantive principles to contemporary conditions.

On the other hand, there are developments in modern logic and experimental science which at least suggest the desirability of new thinking in many areas of the law. We can only hope that as lawyers and judges are freed from some of the demands of routine and drudgery in the performance of their functions, they may have greater opportunity and impulse to devote themselves to the effort of intellectual creation necessary to keep the substance of the law as modern as its techniques promise to become.

Lee Loevinger
Minnesota Supreme Court

c. THE MINICARD SYSTEM FOR HIGH SPEED DATA RETRIEVAL

The MINICARD System is a high-speed information handling system conceived by Eastman Kodak Company and marketed by Recordak Corporation, Kodak's microfilm and office equipment subsidiary. The heart of the MINICARD System is the MINICARD film record on which information is photographically recorded to provide a "permanent" memory. (Fig. 1) This unit record is a tiny piece of microfilm, measuring 16 x 32mm. which provides for compact storage and automatic retrieval. The MINICARD film record has the advantage of coupling the document image and the index code as a single unit. It is this combination that both reduces the space required for storage and enables the high-speed searching which is characteristic of the system. Up to 12 documents--printed pages, charts, tables, photographs or other graphic material can be recorded at 60:1 reduction on each MINICARD film record. In addition to 12 images, the film also may contain 49 characters of code for indexing. By reducing the number of images, the amount of code on each unit record can be increased. For example, a film record with six images has space for 259 characters of code, and a full coded film record (no images) allows up to 455 characters. A binary code system is used which consists of a series of clear and opaque squares exposed on the film.

The MINICARD film records are handled on sticks, each stick holding a total of 2,000 film records. The film records contained on a stick are stored in magazines. There are 50 magazines in each file block, each magazine capable of storing one full stick of film records. The file blocks are stored in a 15-inch by 30-inch by 50-inch, three-drawer file cabinet whose drawers can each hold three file blocks. Thus, each file cabinet can store a total of 900,000 film records, which can contain a maximum of 10,800,000 pages of documents along with 44,000,000 index code characters. By comparison, this amount of original material would occupy over 500 cabinets requiring 2,500 square feet when stored in standard files.

In order to achieve the photographic quality necessary to the success of the MINICARD System, a new film emulsion was created by Eastman Kodak. This fine-grain emulsion is of extremely high resolution--actual resolution of the raw film is more than 4 times greater than standard microfilm or over 1,000 lines per millimeter. Also, new lenses of high resolving power and advanced film handling techniques were developed to fulfill the requirements of the system.

FIGURE 1

Approximately 14 photographic, electronic, mechanical and optical devices are required to record, code, store, sort, select, process and correlate information in the MINICARD System.

Kodak has built four complete MINICARD Systems to date and has two Government contracts for additional systems.

As the first step in setting up a file, the material must be indexed according to pre-established classifications. Cross-indexing of individual items can be extended as far as desired, since duplicate MINICARD film records can be automatically created in the MINICARD equipment for cross-filing. The document code is converted to a punched paper tape, read electronically and recorded on the film. The related graphic material is then recorded. After the roll of film is processed in the MINICARD continuous processor and inspected, it is automatically cut into individual records and stacked on the handling sticks. These "first generation" negatives make up a master file from which duplicates are made for use in expanded working files. A typical document might be recorded and stored in 5 to 10 places, enough to give sufficient complete and separate records to be filed according to as many different concepts as are mentioned in the document. These "second generation" MINICARD film records are sorted into magazines in the storage blocks as directed by the filing code which has been automatically recorded in the proper position on each expanded unit record. Therefore, since each file category contains a film record of every document related to it, only a small section of the files needs to be searched to locate specific information. (Fig. 2 - Input)

Retrieval is accomplished in the following manner. Requests for information are coded and punched into tape. The working file category likely to contain this information is inserted into the selector. Using the tape and a control panel plugboard which indicates selection logic, an electronic selector scans the coded section of each MINICARD film record at a rate of 1200 cards per minute, selecting all those answering the request. A typical search usually takes about 5 minutes.

One pass under the scanner reads the whole record. The equipment is designed to handle multiple record groups as if they were a single record permitting virtually unlimited capacity for associated document images and code.

The selected records are collected in a separate magazine, duplicated automatically to produce "third generation" negatives

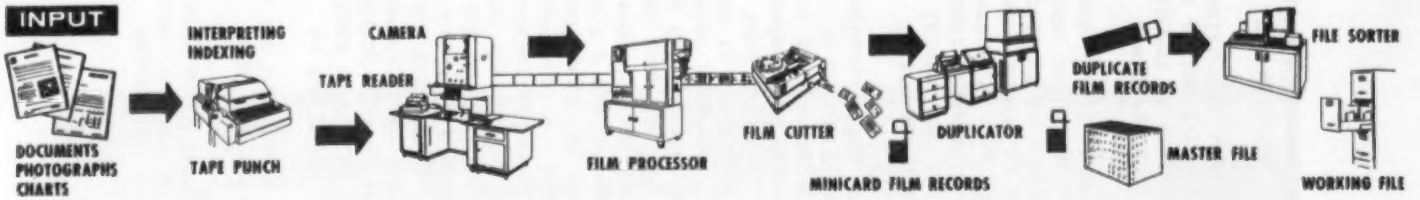


FIGURE 2

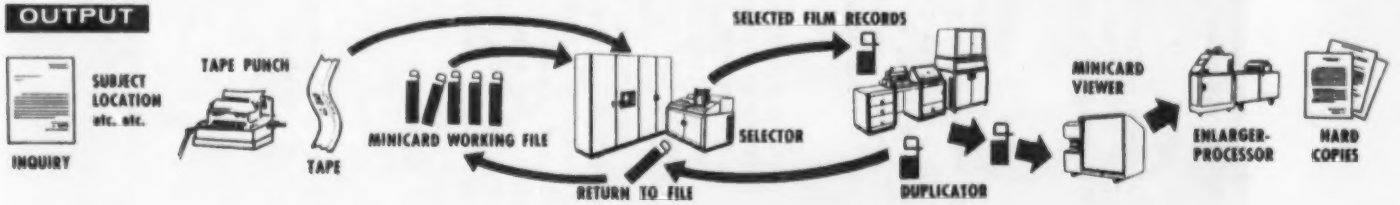


FIGURE 3

and then returned to the file. The "third generation" MINICARD film records go to the requestor who can study them in a viewer or get enlarged hard copy reproductions. In this manner the working file is always intact except for the few moments while the material is being searched. The "third generation" film records may be mounted on aperture cards to be kept as a personal file by the requestor or retained in original form on small handling sticks. These duplicates or the hard copy reproductions can be retained by the requestor in his own files or thrown away when he no longer needs them. (Fig. 3 - Output)

As yet, an off-the-shelf price tag has not been established for individual components or for a complete installation. To establish a price for a particular application, it would be necessary to perform a systems study. If desired, the initial expense of a system could be split among several users. One way to do this would be to set up a centralized operation to handle the job of coding and preparing records. Duplicate film records could then be stored in several working or satellite locations. These might be associated with the central operation or might be separate organizations who would buy the service from the central user. Each of the satellite locations would then need only the searching system.

Summarizing MINICARD System characteristics, we find:

1. The MINICARD film record contains both graphic material and the digital code associated with that material--a true complete unit record.
2. The MINICARD film record combines the mobility of the tabulating card and the compactness of microfilm.
3. Any kind of documentary information will reproduce on MINICARD film records.
4. Single search retrieval at relatively high speeds provides information which is immediately available for reference or for enlarging as original size paper copies.
5. Cross-indexing classification is unlimited.
6. Expendable duplicate film records are delivered thus keeping the working file intact.

The MINICARD System is a modern, electronic microfilm medium for storage and retrieval of documentary information.

Allen M. Rahm
Recordak Corporation

d. STORAGE AND RETRIEVAL OF LEGAL INFORMATION: POSSIBILITIES OF AUTOMATION

The following is a brief summary of portions of a study undertaken at Yale Law School under the sponsorship of the Walter E. Meyer Research Institute of Law. The first part of the study was aimed at making a useful estimate of the magnitude of the legal information-processing problem in the near and long-range future. The second part attempts a first approximation of a comprehensive description of the legal research process. The goal of this description is

1. to begin to distinguish (a) those parts of the legal research process which are routine enough to be amenable to automatic processing from (b) those parts which will continue to require human judgment because of their subtlety or complexity and

2. to give an over-all picture of the legal research process into which the detailed descriptions of the various parts of the total process can be fitted, as these descriptions are produced in later studies.

The first step in determining trends of growth in legal literature is an investigation of what is being referred to by the term "legal literature". The most comprehensive definition of "legal literature" would include as legal literature all documents containing information which could be useful to some lawyer in his professional work. However, such a definition is so broad that it might well include just about all of the world's written matter as legal literature. Perhaps, when the costs of storing information have been sufficiently reduced, the ideal law library will contain all information that has been written. But such is certainly not the case in the year 1960.

However, even a bit of reflection shows that any definition of legal literature is arbitrary, if it sets up a dichotomy between that literature which is and that which is not legal literature. It might seem possible that, as an alternative to a dichotomous definition, all documents could be arrayed along a continuum, so that each document would be associated with a number indicating the degree to which that document is legal literature. In such an array, statutes, case opinions, rulings of administrative agencies and "legal" periodicals would tend to appear at the "high" end of the continuum; items such as arithmetic primers and cookbooks would tend to collect at the

"low" end. But a vast amount of material would inevitably lie in the broad middle of the continuum. If a "definition" is insisted upon, then once such an ordering is made, a point on the continuum could be selected--arbitrarily--and only those documents that are above that point "defined" as legal literature. Who is qualified to say where that point is? Moreover, even the very ordering of documents along such a continuum would be arbitrary. Who is qualified to determine this ordering? Perhaps an appreciation of some of the difficulties involved in attempting to define "legal literature" will make it easier to understand why this problem has not been resolved (or even extensively pursued) in this present study.

In light of the difficulties of specifying a workable definition of "legal literature", the course taken was to look into a few areas which are undoubtedly part of legal literature in order to see what has happened to the rate of growth in these areas through time. The rates of production of the sampled statutes, case opinions, and legal periodicals are summarized in Figures 1-3. It should be emphasized that in Figures 1-3, it is rates of production that are represented, and not the accumulated totals of what has been produced. Figure 4, however, shows the total number of volumes in the Law Library of the U.S. Library of Congress at various times from 1832 to recent years.

Even a cursory examination of these four graphs suggests that if present trends continue, the information-processing problem of lawyers in 1985 is going to be of no small magnitude. Perhaps when lawyers begin to be swamped by a deluge of literature, they will come to a keener appreciation of the efforts being made currently by researchers and librarians in the natural sciences to devise methods for performing some parts of the research process automatically.

Closer examination of the four graphs reveals important differences in the rates of increase in the rates of production of statutes and periodicals as contrasted with that of case opinions. For statutes and periodicals, the rates of increase are increasing, whereas the rates of increase for case opinions are more nearly constant. It is interesting to speculate whether the rate of production of legal periodicals will continue to follow the pattern of change for periodicals in other fields. For, notice that there seems to be a twenty-year lag between the biology curve and the law curve, and a forty-year lag between the medicine curve and the law curve, but that the curves show close correspondences in shape. (There are growth curves in

Figure 1

FEDERAL STATUES IN FORCE IN THE U.S.
1878-1952

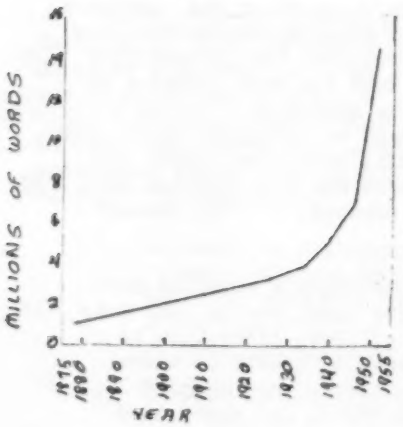


Figure 2

TREND LINES ON AMOUNT OF TEXT IN APPELLATE COURT DECISIONS.
1888-1958

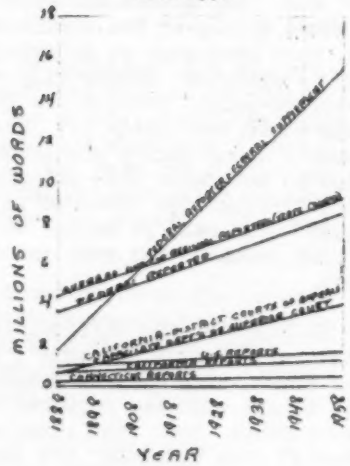


Figure 3

INCREASE IN NUMBER OF PERIODICALS IN BIOLOGY, MEDICINE, AND LAW.
1840-1958

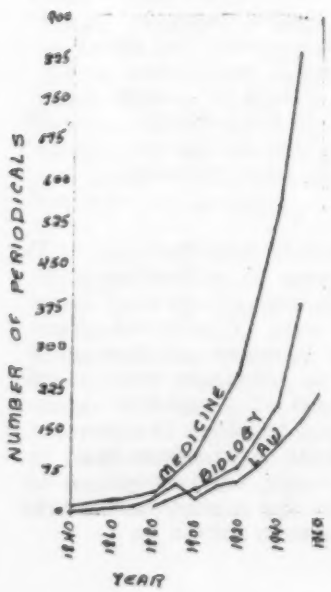
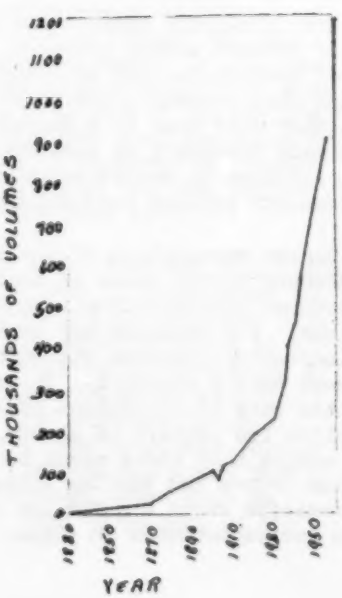


Figure 4

ACCUMULATION OF VOLUMES IN THE LAW LIBRARY OF THE U.S. CONGRESS.
1832-1958



agriculture, in engineering and technology, and in the natural sciences similar to those shown for medicine and biology (R1 at 250).)

Even if the trends toward increased production of legal literature during the past century do not continue during the next twenty-five years--even if the rates of increase level off completely--there will, nevertheless, continue to be a build-up in the total accumulation of legal literature.

There is, of course, the question of literature's becoming obsolete, and such obsolesence will tend to offset this build-up of legal literature to be searched. But this question of obsolesence is another interesting and unresolved one. Just when is a case or a legal article sufficiently old that it is safe to regard it as being obsolete--safe to remove it from the "memory" of the computer so that it will no longer be available to be retrieved? The difficulties with adopting any arbitrary criteria of obsolesence for use in an operating legal information-retrieval system will be apparent to every lawyer. Who knows with such certainty what information is definitely not going to be helpful in advising a client, drafting a document, or persuading a judge?

The pattern of growth of legal literature helps to show why problems of automatic retrieval of information should be of interest to lawyers. If computing machines are to be made useful in legal research, what is needed next is a detailed description of the process of storage and retrieval of legal information as now performed. Only when such a description is available will the computer specialists be in an advantageous position (1) to advise concerning which parts of the present research process can be automated or (2) to suggest entirely new ways of achieving the present goals of legal research.

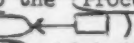
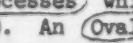
In describing the process of storing and retrieving legal information it is convenient to make use of diagrams, proceeding from a general description to a successively more and more detailed description. Figure 5 represents our most general description of storing and retrieving legal information. A [BOX] is to be regarded as an input to the (Processes) which it precedes (i.e.,  and) and is to be regarded as an output from the (Processes) which it succeeds (i.e.,  and). An (Oval) represents a (Process) for which the [BOXES] are inputs or outputs. Thus, in Figure 1 there is A NEED FOR SOME INFORMATION AS GUIDANCE FOR TAKING SOME ACTION TO HANDLE SOME SITUATION (abbreviated NEED) which is an input to the Research Process. Out of the Research Process

Figure 5

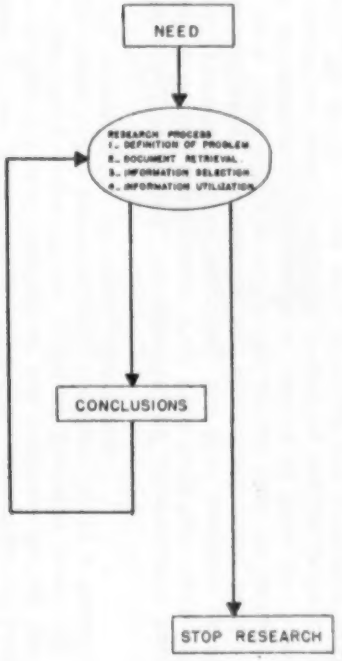
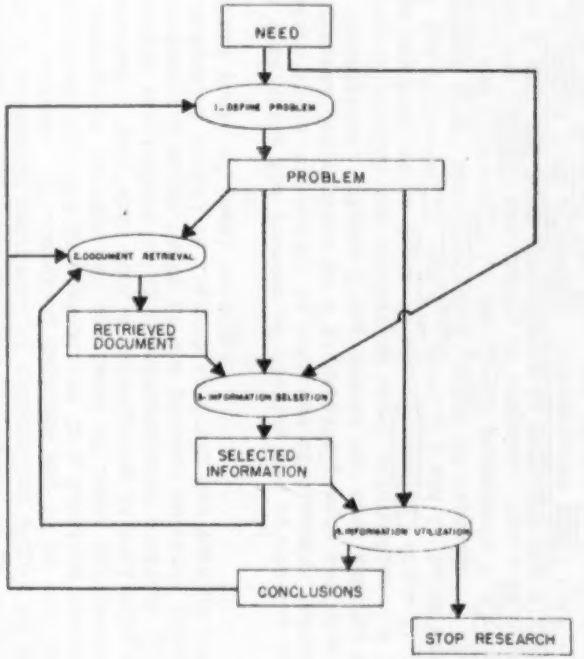


Figure 6



STORAGE AND RETRIEVAL OF LEGAL INFORMATION : AUTOMATIC AND MANUAL

there are two kinds of conclusions as outputs:

--A CONCLUSION TO STOP RESEARCH (abbreviated STOP RESEARCH)
--OTHER CONCLUSIONS (abbreviated CONCLUSIONS).

The CONCLUSIONS feed back as inputs to the Research Process. The description of the storing and retrieving of legal information given in Figure 5 is so general that it is not very useful. It does not indicate which of the various subprocesses of the Research Process are routine enough to be automated.

One step toward a more detailed description is made in Figure 6, in which the relationships between the various subprocesses and their inputs and outputs are shown, as well as the relationships between the subprocesses and the inputs and outputs to the Research Process. The 0-1 NEED is an input to the 1. Definition of Problem Process, which has as its output a 1-2 STATEMENT OF THE PROBLEM (abbreviated PROBLEM). The 1. Definition of Problem Process with its inputs and outputs is the part of the Research Process in which the lawyer clarifies the goals that he is seeking in a particular situation, by considering (either implicitly or explicitly) the discrepancy between the current state of affairs and a preferred future state of affairs.

The 1-2 PROBLEM is an input to the 2. Document Retrieval Process, which has as its output a 2-3 RETRIEVED DOCUMENT. The 2. Document Retrieval Process along with its inputs and outputs is the part of the Research Process in which the lawyer finds the documents that he will read in handling a given situation.

The 2-3 RETRIEVED DOCUMENT, the 0-3 NEED and the 1-3 PROBLEM are inputs to the 3. Information Selection Process, which has as its output 3-4 SELECTED INFORMATION. The 3. Information Selection Process along with its inputs and outputs is the part of the Research Process in which the lawyer reads the documents that have been found and selects such information as

--citations to other relevant documents,
--information contributing to the solution of the problem,
--information contributing to a reformulation of the request and
--information contributing to a redefinition of the problem.

The 3-4 SELECTED INFORMATION and the 1-4 PROBLEM are inputs to the 4. Information Utilization Process, which has as its outputs 4-1 CONCLUSIONS, 4-2 CONCLUSIONS and 4-0 STOP RESEARCH. The 4. Information Utilization Process along with its inputs and outputs is the part of the Research Process in which the lawyer uses the gathered information to choose what action to take in dealing

with the situation and to determine whether to and how to proceed in further research.

Although certain distinctions are beginning to emerge which distinguish processes that can be automated from those that cannot (e.g., one might suspect that parts of 2. Document Retrieval can be done automatically), the description given in Figure 6 is still not sufficiently detailed to be of much assistance. It also has several puzzling features:

--Why is 0-3 NEED an input to 3. Information Selection?
--Why is 3-2 SELECTED INFORMATION an input to 2. Document Retrieval?
--Why is 4-1 CONCLUSION an input to 1. Definition of Problem?
--Why is 4-2 CONCLUSIONS an input to 2. Document Retrieval?

The explanation of these features is facilitated by a more detailed description. Notice that, in moving from a general to a more detailed description of the storing and retrieving of legal information, in Figure 6 a component part of Figure 5 was analyzed into its subcomponents. A still more detailed description of the storage-and-retrieval process results when the following parts of Figure 6 are analyzed into their subcomponents:

2. Document Retrieval
3-4 SELECTED INFORMATION
4. Information Utilization
4-1 CONCLUSION
4-2 CONCLUSIONS

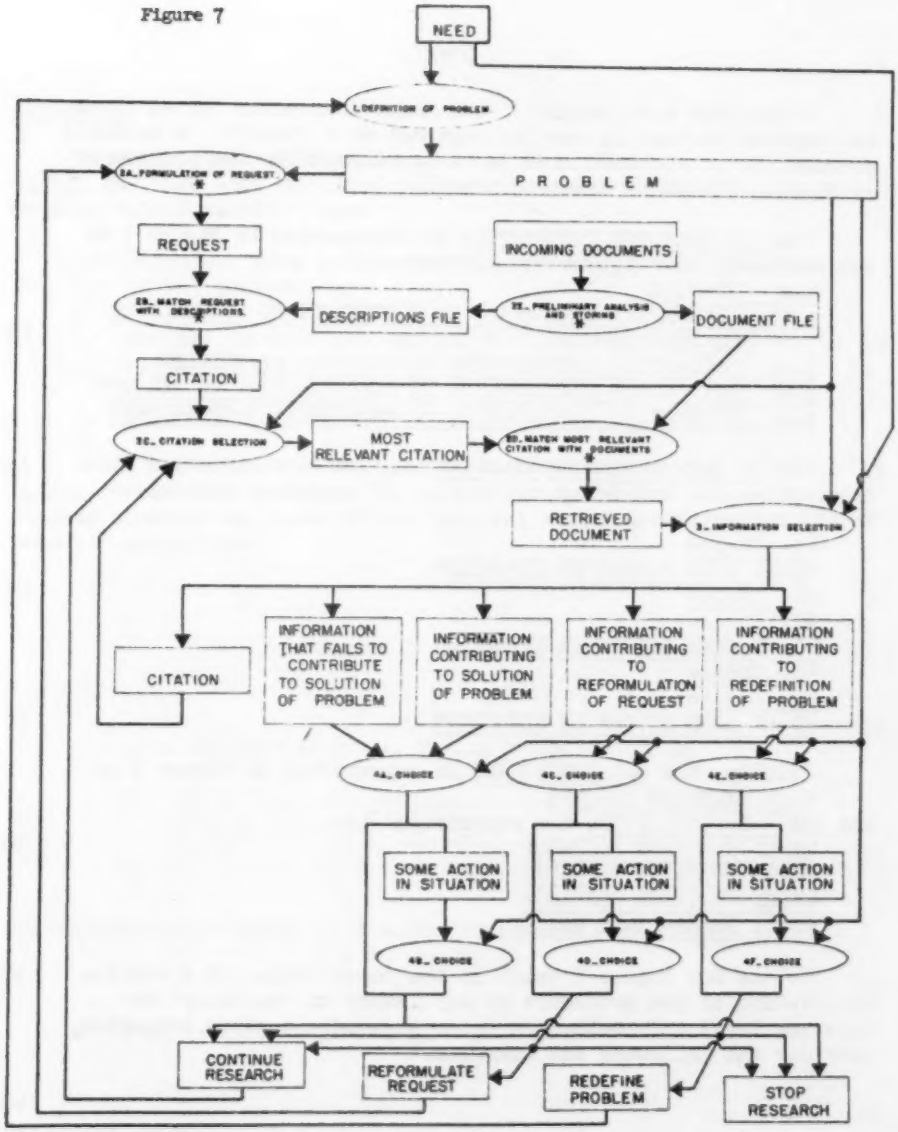
This more detailed description is given in Figure 7.

The 2. Document Retrieval Process is represented in Figure 7 by a combination of the following processes and their inputs and outputs:

0-2E INCOMING DOCUMENTS
2E Preliminary Analysis and Storing
2E-2B DESCRIPTION FILE
2E-2D DOCUMENT FILE

2A. Formulation of Request
2A-2B REQUEST
2E-2B DESCRIPTION FILE
2B. Match Request With Descriptions
2B-2C CITATION
2C. Citation Selection
2C-2D MOST RELEVANT CITATION
2E-2D DOCUMENT FILE
2D. Match Most Relevant Citation With Documents

Figure 7



STORAGE AND RETRIEVAL OF LEGAL INFORMATION : AUTOMATIC AND MANUAL

CHARLES TOUTH HADDAD

A suitable combination of the first four parts listed above corresponds to what is usually regarded as a library. A suitable combination of the last nine parts corresponds to the process of retrieving documents from a library.

The 3-4 SELECTED INFORMATION is represented in Figure 7 by the following five inputs and outputs:

3-2C CITATION
3-4A INFORMATION THAT FAILS TO CONTRIBUTE TO SOLUTION OF THE PROBLEM
3-4A INFORMATION CONTRIBUTING TO SOLUTION OF PROBLEM
3-4C INFORMATION CONTRIBUTING TO REFORMULATION OF REQUEST
3-4E INFORMATION CONTRIBUTING TO REDEFINITION OF PROBLEM

The 4. Information Utilization Process is represented in Figure 7 by the following combination of processes and their inputs and outputs:

4A. Choice
4A-4B SOME ACTION IN SITUATION
4B. Choice
4C. Choice
4C-4D SOME ACTION IN SITUATION
4D. Choice
4E. Choice
4E-4F SOME ACTION IN SITUATION

Finally, the 4-1 CONCLUSION is represented in Figure 7 by:

4F-1 REDEFINE PROBLEM

and the 4-2 CONCLUSIONS are represented by:

4B-2C CONTINUE RESEARCH
4D-2C CONTINUE RESEARCH
4F-2C CONTINUE RESEARCH
4D-2A REFORMULATE REQUEST

To see how Figure 7 helps in the construction of a concise description of the procedure of any lawyer in "manually" retrieving legal information in order to handle a given situation, consider the following six examples.

Example 1

Attorney Adams defines the problem in a situation he is handling, and then gets the most relevant document he can find. From reading this document, Adams

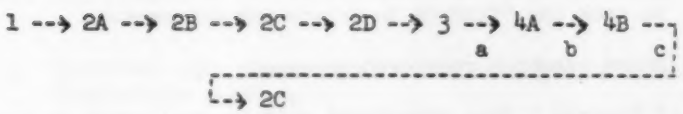
a. discovers some information contributing to the solution of the problem,

b. decides to take some action with respect to the given situation and

c. decides to continue research.

Using the alpha-numeric symbols associated with the processes represented in Figure 7, we can construct the following diagram showing the kinds of and temporal sequence of Adams' research activities.

Diagram 1



In interpreting Diagram 1, consider first the following part:

1 --> 2A --> 2B --> 2C --> 2D --> 3

This part of Diagram 1 is intended to indicate:

1. Adams defines the PROBLEM from the description of the situation that has been presented to him.

2A. Adams formulates a REQUEST for information. This he may do explicitly by external communication to another person (e.g., asking the librarian for all the articles dealing with the INCOME TAX effects of BEQUESTS to CHARITABLE ORGANIZATIONS that have appeared during the past 10 years in HARV. L. REV., COL. L. REV. or YALE L.J.), or he may make a private "request" to himself (e.g., by deciding simply to reach for his copy of the Internal Revenue Code and the Treasury Regulations).

2B. Adams (or his librarian) matches the REQUEST with the descriptions in the DESCRIPTION FILE (e.g., he looks up the word CHARITABLE ORGANIZATIONS in the card catalogue or some other index).

2C. From the list of CITATIONS found by matching REQUEST with DESCRIPTIONS, the CITATION of the document that ought to be read first is obtained. The citation of that document is called the MOST RELEVANT CITATION.

2D. Adams (or his librarian) matches the MOST RELEVANT CITATION with the documents in the DOCUMENT FILE, i.e., he uses the CITATION to locate the document.

3. Adams reads the retrieved document.

The part of Diagram 1 just considered will be duplicated in the diagrams of each of the examples to follow.

The part of Diagram 1 labeled 'a' indicates that in reading the first document retrieved, Adams found some information contributing to the solution of the problem. The part labeled 'b' indicates that on the basis of this information he decided to take some action with respect to the situation. Finally, the part labeled 'c' indicates that Adams concluded that further research was needed, and so returned to the 2C. Citation Selection Process to begin anew the process of obtaining further documents to read.

Example 2

Attorney Baker defines his problem and reads the most relevant document that he can find, but he

a. does not discover any information that contributes to

the solution of his problem (Nevertheless he must stop research because his boss has given him a more important job to do now.),

b. decides nevertheless to take some action with respect to the given situation and

c. decides to stop research.

By using Figure 7 the following diagram can be constructed for the second example.

Diagram 2

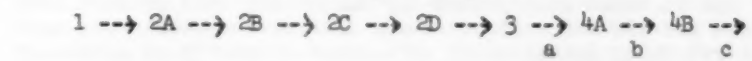


Diagram 2 is just like Diagram 1 except that here 4B feeds to 0, i.e., the research process stops, whereas in Diagram 1 there is feedback to 2C, i.e., the research process continues.

Example 3

. . . Attorney Charles

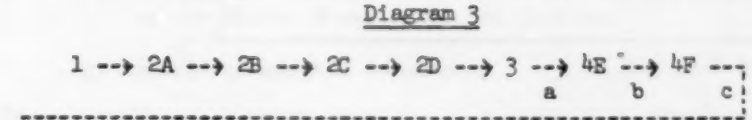
a. discovers some information contributing to a redefinition of the problem,

b. decides to take some action with respect to the given situation and

c. decides to redefine the problem and begin research on the new problem.

By using Figure 7 the following diagram can be constructed for the third example.

Diagram 3



The most significant feature of Diagram 3 is c, the long feedback from 4F all the way back to the beginning of the Research Process at 1. What c indicates is that the output of the 4F Choice Process is a conclusion to redefine the problem, which feeds back to 1. Definition of Problem Process.

As a lawyer acquires information in working on a problem, he often becomes aware that his initial requests for information need to be modified. This may be because the acquired information indicates that the initial definition of the problem is erroneous and in need of modification. In doing research with respect to a given situation, there may well be a series of successive approximations in the defining of the problem or in the formulating of the request. In the present "manual" form of legal research, there is a continual interaction between the lawyer and his "retrieval system". And it seems likely that any adequate system of retrieval must provide for such continual interaction between lawyer and retrieval system--even an automatic one. Such interaction is necessary so that what is asked of the system can be modified as the inflow of information enables the lawyer to clarify his goals.

Example 4

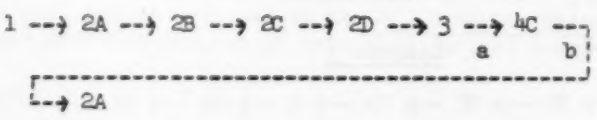
. . . Attorney Daniels

a. discovers information contributing to a reformulation of his initial request and

b. decides to reformulate the request without taking any action with respect to the situation, and to begin retrieving documents anew on the basis of the reformulated request.

By using Figure 7 the following diagram can be constructed for the fourth example.

Diagram 4

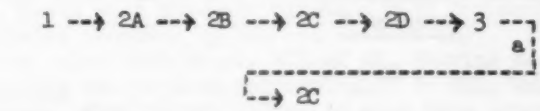


Example 5

. . . Attorney Evans

a. discovers a new citation to a relevant document.

Diagram 5



When Evans is referred back to the 2C. Citation Selection Process, he may choose either to continue reading the same document he had been reading when the new citation was encountered, or to read some other document. It is not inconceivable that the criteria for choosing citations in the 2C. Citation Selection Process could be made sufficiently explicit that some or all of this process could be performed automatically.

Example 6

. . . Attorney Ford

a. discovers a new citation by reading the retrieved document,

b. selects a new most relevant citation,

c. retrieves the document which matches that new most relevant citation,

d. discovers information contributing to redefinition of the problem by reading the retrieved document,

e. decides to take some action with respect to the situation,

f. decides to redefine the problem and begin research anew on the basis of the redefined problem,

g. retrieves another document,

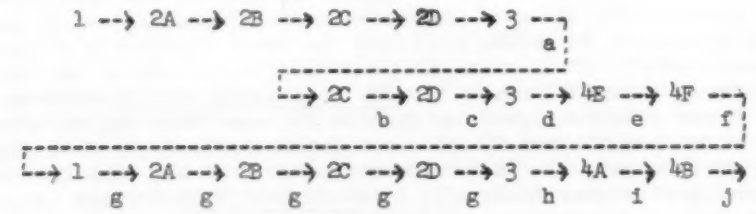
h. discovers information contributing to solution of the problem by reading this most recently retrieved document,

i. decides to take some action with respect to the situation and

j. decides to stop research.

Using Figure 7 it is easy to generate the following concise description of Ford's research procedure.

Diagram 6



The usual pattern in doing research to handle a complicated situation probably involves many more such cycles through the research process -- many redefinitions in order to adequately refine the problem, many reformulations in order to get the request that will produce the most relevant information, perhaps hundreds or even thousands of citations discovered and evaluated, and perhaps many intermediate actions taken to deal with the situation -- before any ultimate solution is determined and research is stopped.

The several puzzling features of Figure 6 may seem somewhat less puzzling now that Figure 7 has been scrutinized closely in tracing through the examples just given. Since one of the outputs of the 3. Information Selection Process is 3-4E INFORMATION CONTRIBUTING TO REDEFINITION OF PROBLEM, there must be criteria for the selection of such information. Certainly, 0-3 NEED will be wanted as an input to 3. Information Selection for inclusion among such criteria. In Figure 7 it also becomes apparent why 3-2 SELECTED INFORMATION (of Figure 6) feeds back to 2. Document Retrieval (of Figure 6); the 3-2C CITATIONS selected in the 3. Information Selection Process are fed back into the 2D. Citation Selection Process. Similarly, the reasons why 4-1 CONCLUSION is an input to 1. Definition of Problem and 4-2 CONCLUSIONS an input to 2. Document Retrieval are evident from an inspection of Figure 7.

The degree of detail in Figure 7 provides a description of the storing and retrieving of legal information that is at last

beginning to look interesting from the viewpoint of distinguishing processes that can be automated from those that require human intervention. It has been mentioned that perhaps some or all of the 2C. Citation Selection Process could be done automatically -- if the criteria for selection could be made sufficiently explicit. The processes labeled with an '*' in Figure 7 are ones that can be done automatically in some retrieval systems, e.g., 2A. Formulation of Request, 2B. Match Request with Descriptions and 2D. Match Most Relevant Citation with Documents are automated in a system proposed by IBM researchers (R2). It should be mentioned also that nearly all of 2E. Storing and Preliminary Analysis can be done automatically in this same IBM retrieval system. What is yet needed to make this process wholly automatic is a general purpose scanner for transliterating prose into a machine-readable form. It should be noted that this IBM retrieval system seems to put much emphasis upon the frequency of occurrence of certain of the words in a document and relatively little emphasis upon the relationships between these words. (See Newman's comment on this with respect to patent-search problems, MULL 60J1a47.)

Observe that it is in the 2. Document Retrieval Process that the parts that seem most susceptible to being automated occur. It is somewhat difficult to conceive how much, if any, of the 1. Definition of Problem Process can at the present time be delegated to a machine. The same comment applies to the 3. Information Selection Process, with one small qualification: recognition of citations can probably be done automatically.

We have presented our first approximation of a comprehensive description of the storing and retrieving of legal information. Hopefully, this description will stimulate comments that will aid in revising it to eliminate errors and omissions. Then, the research task for the future is the identification of subcomponents of the parts distinguished in Figure 7, and only after that can it be ascertained which subcomponents will next surrender to the wizardry of the electronic age.

Layman E. Allen,
Robin B. S. Brooks and
Patricia A. James
Yale Law School

REFERENCES

R1 Hearings before a Subcommittee of the Senate Committee on Government Operations, 85th Cong., 2d Sess., pt. 1 (1958).

R2 See, e.g., Luhn, H. P., "Auto-encoding of Documents for Information Retrieval Purposes", in MODERN TRENDS IN DOCUMENTATION, (Ed. by M. Boaz, 1959).

3. REVIEWS AND ABSTRACTS

a. Machine Data Processing Systems For The Trial Lawyer
by Roy N. Freed, THE PRACTICAL LAWYER,
April 1960, p. 73.

(Author's Abstract)

This article describes a newly devised technique for managing large accumulations of trial evidence in major litigation by means of a punch card machine data processing system. The technique requires a careful classification of the various facets of each case, based in large measure on the underlying trial strategy. Each item of evidence is coded in terms of the categories thus established, for retrieval as the litigation progresses. The greater effort in setting up a system of classification than normally might be expended probably will be rewarded in terms of better case preparation. To facilitate the process, forms are suggested by which the lawyer simultaneously analyzes each item and conveys coding instructions to the card puncher.

As advantages of the new technique, it is claimed that the attention of lawyers to non-legal details will be minimized and that the ready availability of all items of evidence will be assured despite personnel changes. The great expense of trial record searching by lawyers will be reduced, and the failure to discover important facts because of human perception failings should be eliminated.

Because only a card punch and a sorter are required, the expense of such a technique is comparatively low and might be less than savings accomplished by it.

It is suggested that the technique be utilized by courts and administrative tribunals as well as by the litigants as a means of reducing the time and effort required for reaching decisions in large cases. Such a use would be more likely if opposing counsel could agree on an index of the record in terms of which the tribunal's deck of cards could be set up but which would not disclose the trial strategy of any party.

The article attempts to acquaint lawyers with the mechanics of machine data processing in the hope that their awareness of its intricacies and of its substantial benefits will stimulate greater efforts toward the performance of legal research with the aid of a more sophisticated, fully electronic variety of such a machine system.

4. BIBLIOGRAPHY

a. BIBLIOGRAPHY ON DATA RETRIEVAL FOR LEGAL RESEARCH

1. Bush, V.
"As We May Think," 176 ATLANTIC MONTHLY 101 (July 1945).

2. Casey, P.
PUNCHED CARDS (1951)

3. Harper, L. A.
"Legal Research, Technology and the Future," 104 CALIF. S.B.J. (1949).

4. Jackson, J. C.
"Records of Research," 35 J.P.O.S. 239 (1953).

5. Jacobstein, J. M.
"Scientific Aids for Legal Research," 31 CHICAGO KENT LAW REVIEW 236-245 (1953)

6. Kelso, L. O.
"Does the Law Need a Technological Revolution," 18 ROCKY MOUNTAIN LAW REVIEW 378 (1946)

7. Maguire, J. M.
"Adaptation of Modern Mechanical and Electrical Aids to the Assembly of Tax Information," 7 N.Y.U. FED. TAX INST. 1008 (1949).

8. Merryman, J. H.
"Legal Research Without Books," 44 LAW LIB. J. 7 (1951).

9. Perry, J. W., Kent, A., and Berry, M. M.
MACHINE LITERATURE SEARCHING (1956).

10. Taube, Mortimer
STUDIES IN COORDINATE INDEXING ( ).

11. Walkin, P. A.
"Can it Happen Here?" 18 SHINGLE 181 (1955).

SOURCES

1. AMERICAN DOCUMENTATION
Interscience Publishers, Inc., 250 Fifth Ave., New York 1, N.Y. Published for the American Documentation Institute in cooperation with the School of Library Science, Western

Reserve University.

2. DATA PROCESSING DIGEST
Canning Sisson and Associates, 914 South Robertson Boulevard, Los Angeles 35, California.

3. THE EASTMAN MINICARD SYSTEM
Eastman Kodak Company, Rochester 4, New York, 1955.

4. JOURNAL OF LEGAL EDUCATION

5. LAW LIBRARY JOURNAL

6. NEWSLETTER
Center for Documentation and Communication Research, School of Library Science, Western University, Cleveland 6, Ohio.

7. OFFICE OF RESEARCH AND DEVELOPMENT
U. S. Patent Office.

8. SPECIAL LIBRARIES ASSOCIATION

9. ZATOR COMPANY
Boston, Massachusetts.

Richard F. C. Hayden
Chairman, EDR Committee
Member of Los Angeles Bar

b. SELECTED BIBLIOGRAPHY OF SOURCES OF CURRENT INFORMATION ON DOCUMENTATION

1. Council on Library Resources, Inc.
ANNUAL REPORT, 1025 Connecticut Avenue, N.W. Washington 6, D. C.

2. Council on Library Resources, Inc.
RECENT DEVELOPMENTS, 1025 Connecticut Avenue, N.W. Washington 6, D. C.

3. IBM Service Bureau Corporation
LITERATURE ON INFORMATION RETRIEVAL AND MACHINE TRANSLATION. 2nd. edition, June 1959, edited by Peter James. Request from IBM, Service Bureau Corporation, 425 Park Avenue, New York 22, N. Y.

4. International Federation for Documentation
F.I.D. NEWS BULLETIN

5. PROBLEMS IN INFORMATION STORAGE AND RETRIEVAL
Pergamon Press, 122 East 55th Street, New York 22, N. Y.
$15.00 per year for private individuals from Pergamon Press, 122 East 55th Street, New York 22, N. Y.

6. Stanford Research Institute
Division of Engineering Research. BIBLIOGRAPHY ON THE MECHANIZATION OF INFORMATION RETRIEVAL. By Charles P. Bourne, Menlo Park, California. Request.

7. U. S. Department of Health, Education and Welfare
Office of Education. Library Services Branch. LIBRARY RESEARCH IN PROGRESS. Washington 25, D. C. Request.

8. U. S. National Science Foundation
CURRENT RESEARCH AND DEVELOPMENT IN SCIENTIFIC DOCUMENTATION. Washington 25, D. C. Semiannual, October and April. 50¢ per issue from U. S. Superintendent of Documents

9. U. S. National Science Foundation
NONCONVENTIONAL TECHNICAL INFORMATION SYSTEMS IN CURRENT USE. Washington 25, D. C. (No. 1 unavailable) Nos. 2 (30¢) and Supplement to No. 2 (25¢) from U. S. Superintendent of Documents.

10. U. S. National Science Foundation
SCIENTIFIC INFORMATION NOTES (Reporting National and International Developments in Scientific and Technical Information Dissemination) Washington 25, D. C. $1.25 for annual subscription from U. S. Superintendent of Documents.

11. Western Reserve University
School of Library Science. NEWSLETTER, CENTER FOR DOCUMENTATION AND COMMUNICATION RESEARCH. Cleveland 6, Ohio. Request.

Prepared by Devices Subcommittee,
Am. Ass. of Law Lib. Committee
On the Application of Scientific
Devices to Legal Literature

6. LETTERS OF READERS AND OTHER MISCELLANEOUS ITEMS

a. POLICY STATEMENT OF WEST PUBLISHING COMPANY ON AUTOMATIC RETRIEVAL OF LEGAL INFORMATION

In a letter to the editor of MULL, William W. Marvin, Vice President of the West Publishing Company, enclosed the following statement of policy on West's present position with regard to automatic retrieval of information in law.

"What the future may hold in this day of inventive advances, we do not attempt to foretell. Our findings, as of today, convince us that although electronic data processing computors have been developed to an incredible degree, they still do not, as of now, offer a practical medium for legal research.

"There have been several attempts to adapt microfilm to current legal publications in daily use. Here, too, present day developments are not adequate for everyday general library research which is the real goal. We could not ask our subscribers to adopt, at a very substantial cost, a process still in an experimental stage.

"Our responsibility as major publishers obligates us to consider the needs of the profession both collectively and individually and we are not closing our minds to possible future developments. The nature of the legal processes with the many shades of distinction do not lend themselves to adequate classification through automation by any present system.

"Several opportunities to explore the development of a project similar to that you suggest have been proposed to us and we have not been deterred by cost or technical difficulties but rather that our study of electronic processing computors have shown very clearly that they would not be adequate nor acceptable to the legal profession.

"Several bar association committees have made studies and contacted us on the subject. We have advised them that West Publishing Company is continuing its study of all possible methods of improving and shortening legal research but there would be no justification in our offering to the legal profession a plan of mechanical storage and automatic research of the law that employs any present electronic computors, tape recorders, or other mechanical devices."

In his letter Mr. Marvin stated further:

"We assure you that West Publishing Company is giving constant thought and study to the subject and in this day of scientific developments, anything may be possible for the future; but today there is no development that offers a practical plan that would have the support and acceptance of the legal profession."

b. PROGRAM OF THE EDR COMMITTEE AT THE 1960 ANNUAL MEETING OF THE ABA IN WASHINGTON, D.C.

1. Demonstration of the Use of Electronic Data Processing Systems in Legal Research

Speaker: John F. Horty, Director, Health Law Center, University of Pittsburgh, and others

Demonstrations of the application of electronic data processing systems to legal research will be presented under the joint sponsorship of the EDR Committee and the University of Pittsburgh.

2. Demonstration of the Use of Electronic Data Processing Equipment in Patent Office Searching

Speaker: Donald D. Andrews, Director, Research and Development, United States Patent Office, and others

Demonstrations of the application of electronic data processing equipment to Patent Office searching will be presented under the joint sponsorship of the Electronic Data Retrieval Committee and the United States Patent Office.

Both of the foregoing demonstrations will be made periodically every day during the Convention, commencing Saturday, August 27. Because of the limited space available at IBM headquarters, it will be necessary to obtain tickets for reservations for specific demonstrations at the Registration Desk at the Hotel Statler. Bus transportation will be available to the site of the demonstrations at IBM headquarters at 1220 Nineteenth Street, N.W., in Washington, D.C.

3. Meeting of the EDR Committee, Monday, Aug. 29, Congressional Room, Hotel Statler, 9:00 AM to 12:00 AM.

9:00 INTRODUCTION — Cecil E. Burney, Chairman of the Bar Activities Section, Member of Texas Bar, Corpus Christi, Texas

GENERAL REMARKS — Richard F.C. Hayden, Chairman of Electronic Data Retrieval Committee, Member of California Bar, Los Angeles, California

9:10 "Experience with the Application of Electronic Data Processing Systems in General Law"
John F. Horty, Member of Pennsylvania Bar, Director, Health Law Center, University of Pittsburgh.

9:35 "Experience with Electronic Searching of United States Patents"
Donald D. Andrews, Member of District of Columbia Bar, Director of Research and Development, United States Patent Office.

10:00 PANEL DISCUSSION — "Progress and Problems in the application of Electronic Data Processing Systems to Legal Research"

Moderator: Reed C. Lawlor, Member of California Bar, Los Angeles, California

Panelists: Vincent P. Biunno, Member of New Jersey Bar, Newark, New Jersey

Vincent Fiordalisi, Professor of Law and Law Librarian, Rutgers University

Allen Kent, Associate Director, Center for Documentation and Communication Research, Western Reserve University

H. Peter Luhn, Advisor, Business Intelligence Systems, Advanced Systems Development Section, International Business Machines Corporation, Poughkeepsie, New York

The foregoing panelists will discuss legal research methods in general, the capabilities of data processing systems for legal research, the history of the application of mechanized techniques to legal research, and the economics and advantages of mechanized legal research.

Reed C. Lawlor
Chairman of the 1960 Program
of the EDR Committee

c. PRACTICING LAW INSTITUTE COURSE DEALING WITH INFORMATION RETRIEVAL

The Practicing Law Institute (20 Vesey St., New York 7, N.Y.) has announced a course on current problems in patent law. The session scheduled for 7:30 PM, Monday, July 11, at the Statler-Hilton Hotel in New York City, is entitled "Information Retrieval". It is described as follows:

> "Storing and retrieving information as mechanical aids in searching; the patent lawyer's interest; practical aspects and ultimate objectives."

University of Michigan
Law Library
Ann Arbor, Michigan
628

Citations to MULL

To keep citations brief, the following system will be used in making citations to MULL. Example: MULL 60M6c

60 Volume number -- coincides with year of publication
M Month of publication -- M March S September
J June D December
6 Section number
c Subsection number
Page number (if necessary)

The first issue of MULL appeared in September 1959.

Since one of the major purposes of MULL is to provide a vehicle for the exchange of information about

1. the uses of modern logic in law and

2. modern methods of information retrieval in law

readers are urged to send in news about research in progress that will be appropriate for inclusion in MULL.